



OICE

## RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/746,660A

TIME: 11:32:37

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

3 <110> APPLICANT: Pompejus, Markus  
 4 Kroger, Burkhard  
 5 Schroder, Hartwig  
 6 Zelder, Oskar  
 7 Haberhauer, Gregor  
 8 Kim, Jun-Won  
 9 Lee, Heung-Schick  
 10 Hwang, Byung-Joon  
 12 <120> TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
 13 METABOLIC PATHWAY PROTEINS  
 15 <130> FILE REFERENCE: BGI-121CP2  
 17 <140> CURRENT APPLICATION NUMBER: US 09/746,660A  
 18 <141> CURRENT FILING DATE: 2000-12-22  
 20 <150> PRIOR APPLICATION NUMBER: 09/606740  
 21 <151> PRIOR FILING DATE: 2000-06-23  
 23 <150> PRIOR APPLICATION NUMBER: 09/603124  
 24 <151> PRIOR FILING DATE: 2000-06-23  
 26 <150> PRIOR APPLICATION NUMBER: 60/141031  
 27 <151> PRIOR FILING DATE: 1999-06-25  
 29 <150> PRIOR APPLICATION NUMBER: 60/142101  
 30 <151> PRIOR FILING DATE: 1999-07-02  
 32 <150> PRIOR APPLICATION NUMBER: 60/148613  
 33 <151> PRIOR FILING DATE: 1999-08-12  
 35 <150> PRIOR APPLICATION NUMBER: 60/187970  
 36 <151> PRIOR FILING DATE: 2000-03-09  
 38 <150> PRIOR APPLICATION NUMBER: DE 19931420.9  
 39 <151> PRIOR FILING DATE: 1999-07-08  
 41 <160> NUMBER OF SEQ ID NOS: 125  
 43 <170> SOFTWARE: PatentIn Vers. 2.0  
 45 <210> SEQ ID NO: 1  
 46 <211> LENGTH: 1840  
 47 <212> TYPE: DNA  
 48 <213> ORGANISM: Corynebacterium glutamicum  
 50 <220> FEATURE:  
 51 <221> NAME/KEY: CDS  
 52 <222> LOCATION: (363)..(1676)  
 54 <400> SEQUENCE: 1  
 55 cagaaactgt gtgcagaaat gcatgcagaa aaaggaaagt tcgggcccaag atgggtgttt 60  
 57 ctgtatgccg atgacggat ctttgacagc tgggtatgcg acaaatacc gagagttgtt 120  
 59 aattcttaac aatggaaaag taacattgag agatgattta taccatcctg caccatttag 180  
 61 agtggggcta gtcatacccc cataacccta gctgtaogca atcgatttca aatcagttgg 240  
 63 aaaaagtcaa gaaaattacc cgagaattaa ttataccac acagtctatt gcaatagacc 300  
 65 aagctgttca gtaggggtgca tgggagaaga atttctaata aaaaactctt aaggacctcc 360

85  
**ENTERED**

## RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/746,660A

TIME: 11:32:37

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

```

67 aa atg cca aag tac gac aat tcc aat gct gac cag tgg ggc ttt gaa      407
68   Met Pro Lys Tyr Asp Asn Ser Asn Ala Asp Gln Trp Gly Phe Glu
69     1           5           10           15
71 acc cgc tcc att cac gca ggc cag tca gta gac gca cag acc agc gca      455
72 Thr Arg Ser Ile His Ala Gly Gln Ser Val Asp Ala Gln Thr Ser Ala
73           20           25           30
75 cga aac ctt ccg atc tac caa tcc acc gct ttc gtg ttc gac tcc gct      503
76 Arg Asn Leu Pro Ile Tyr Gln Ser Thr Ala Phe Val Phe Asp Ser Ala
77           35           40           45
79 gag cac gcc aag cag cgt ttc gca ctt gag gat cta ggc cct gtt tac      551
80 Glu His Ala Lys Gln Arg Phe Ala Leu Glu Asp Leu Gly Pro Val Tyr
81           50           55           60
83 tcc cgc ctc acc aac cca acc gtt gag gct ttg gaa aac cgc atc gct      599
84 Ser Arg Leu Thr Asn Pro Thr Val Glu Ala Leu Glu Asn Arg Ile Ala
85           65           70           75
87 tcc ctc gaa ggt ggc gtc cac gct gta gcg ttc tcc tcc gga cag gcc      647
88 Ser Leu Glu Gly Gly Val His Ala Val Ala Phe Ser Ser Gly Gln Ala
89 80           85           90           95
91 gca acc acc aac gcc att ttg aac ctg gca gga gcg ggc gac cac atc      695
92 Ala Thr Thr Asn Ala Ile Leu Asn Leu Ala Gly Ala Gly Asp His Ile
93           100          105          110
95 gtc acc tcc cca cgc ctc tac ggt ggc acc gag act cta ttc ctt atc      743
96 Val Thr Ser Pro Arg Leu Tyr Gly Gly Thr Glu Thr Leu Phe Leu Ile
97           115          120          125
99 act ctt aac cgc ctg ggt atc gat gtt tcc ttc gtg gaa aac ccc gac      791
100 Thr Leu Asn Arg Leu Gly Ile Asp Val Ser Phe Val Glu Asn Pro Asp
101           130          135          140
103 gac cct gag tcc tgg cag gca gcc gtt cag cca aac acc aaa gca ttc      839
104 Asp Pro Glu Ser Trp Gln Ala Ala Val Gln Pro Asn Thr Lys Ala Phe
105           145          150          155
107 ttc ggc gag act ttc gcc aac cca cag gca gac gtc ctg gat att cct      887
108 Phe Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro
109 160           165          170          175
111 gcg gtg gct gaa gtt gcg cac cgc aac agc gtt cca ctg atc atc gac      935
112 Ala Val Ala Glu Val Ala His Arg Asn Ser Val Pro Leu Ile Ile Asp
113           180          185          190
115 aac acc atc gct acc gca gcg ctc gtg cgc ccg ctc gag ctc ggc gca      983
116 Asn Thr Ile Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala
117           195          200          205
119 gac gtt gtc gtc gct tcc ctc acc aag ttc tac acc ggc aac ggc tcc      1031
120 Asp Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser
121           210          215          220
123 gga ctg ggc ggc gtg ctt atc gac ggc gga aag ttc gat tgg act gtc      1079
124 Gly Leu Gly Gly Val Leu Ile Asp Gly Gly Lys Phe Asp Trp Thr Val
125           225          230          235
127 gaa aag gat gga aag cca gta ttc ccc tac ttc gtc act cca gat gct      1127
128 Glu Lys Asp Gly Lys Pro Val Phe Pro Tyr Phe Val Thr Pro Asp Ala
129 240           245          250          255
131 gct tac cac gga ttg aag tac gca gac ctt ggt gca cca gcc ttc ggc      1175

```

## RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/746,660A

TIME: 11:32:37

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

```

132 Ala Tyr His Gly Leu Lys Tyr Ala Asp Leu Gly Ala Pro Ala Phe Gly
133          260          265          270
135 ctc aag gtt cgc gtt ggc ctt cta cgc gac acc ggc tcc acc ctc tcc 1223
136 Leu Lys Val Arg Val Gly Leu Leu Arg Asp Thr Gly Ser Thr Leu Ser
137          275          280          285
139 gca ttc aac gca tgg gct gca gtc cag ggc atc gac acc ctt tcc ctg 1271
140 Ala Phe Asn Ala Trp Ala Ala Val Gln Gly Ile Asp Thr Leu Ser Leu
141          290          295          300
143 cgc ctg gag cgc cac aac gaa aac gcc atc aag gtt gca gaa ttc ctc 1319
144 Arg Leu Glu Arg His Asn Glu Asn Ala Ile Lys Val Ala Glu Phe Leu
145          305          310          315
147 aac aac cac gag aag gtg gaa aag gtt aac ttc gca ggc ctg aag gat 1367
148 Asn Asn His Glu Lys Val Glu Lys Val Asn Phe Ala Gly Leu Lys Asp
149 320          325          330          335
151 tcc cct tgg tac gca acc aag gaa aag ctt ggc ctg aag tac acc ggc 1415
152 Ser Pro Trp Tyr Ala Thr Lys Glu Lys Leu Gly Leu Lys Tyr Thr Gly
153          340          345          350
155 tcc gtt ctc acc ttc gag atc aag ggc ggc aag gat gag gct tgg gca 1463
156 Ser Val Leu Thr Phe Glu Ile Lys Gly Gly Lys Asp Glu Ala Trp Ala
157          355          360          365
159 ttt atc gac gcc ctg aag cta cac tcc aac ctt gca aac atc ggc gat 1511
160 Phe Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Ile Gly Asp
161          370          375          380
163 gtt cgc tcc ctc gtt gtt cac cca gca acc acc acc cat tca cag tcc 1559
164 Val Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser
165          385          390          395
167 gac gaa gct ggc ctg gca cgc gcg ggc gtt acc cag tcc acc gtc cgc 1607
168 Asp Glu Ala Gly Leu Ala Arg Ala Gly Val Thr Gln Ser Thr Val Arg
169 400          405          410          415
171 ctg tcc gtt ggc atc gag acc att gat gat atc atc gct gac ctc gaa 1655
172 Leu Ser Val Gly Ile Glu Thr Ile Asp Asp Ile Ile Ala Asp Leu Glu
173          420          425          430
175 ggc ggc ttt gct gca atc tag ctttaaatag actcacccca gtgcttaaag 1706
176 Gly Gly Phe Ala Ala Ile
177          435
179 cgctgggttt ttctttttca gactcgtgag aatgcaaact agactagaca gagctgtcca 1766
181 tatacactgg acgaagtttt agtcttgtcc acccagaaca ggcggttatt ttcatgcccc 1826
183 cctcgcgcc ttca 1840
186 <210> SEQ ID NO: 2
187 <211> LENGTH: 437
188 <212> TYPE: PRT
189 <213> ORGANISM: Corynebacterium glutamicum
191 <400> SEQUENCE: 2
192 Met Pro Lys Tyr Asp Asn Ser Asn Ala Asp Gln Trp Gly Phe Glu Thr
193 1 5 10 15
195 Arg Ser Ile His Ala Gly Gln Ser Val Asp Ala Gln Thr Ser Ala Arg
196 20 25 30
198 Asn Leu Pro Ile Tyr Gln Ser Thr Ala Phe Val Phe Asp Ser Ala Glu
199 35 40 45

```

## RAW SEQUENCE LISTING

DATE: 02/12/2002

PATENT APPLICATION: US/09/746,660A

TIME: 11:32:37

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

```

201 His Ala Lys Gln Arg Phe Ala Leu Glu Asp Leu Gly Pro Val Tyr Ser
202      50      55      60
204 Arg Leu Thr Asn Pro Thr Val Glu Ala Leu Glu Asn Arg Ile Ala Ser
205 65      70      75      80
207 Leu Glu Gly Gly Val His Ala Val Ala Phe Ser Ser Gly Gln Ala Ala
208      85      90      95
210 Thr Thr Asn Ala Ile Leu Asn Leu Ala Gly Ala Gly Asp His Ile Val
211      100      105      110
213 Thr Ser Pro Arg Leu Tyr Gly Gly Thr Glu Thr Leu Phe Leu Ile Thr
214      115      120      125
216 Leu Asn Arg Leu Gly Ile Asp Val Ser Phe Val Glu Asn Pro Asp Asp
217      130      135      140
219 Pro Glu Ser Trp Gln Ala Ala Val Gln Pro Asn Thr Lys Ala Phe Phe
220 145      150      155      160
222 Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro Ala
223      165      170      175
225 Val Ala Glu Val Ala His Arg Asn Ser Val Pro Leu Ile Ile Asp Asn
226      180      185      190
228 Thr Ile Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala Asp
229      195      200      205
231 Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser Gly
232      210      215      220
234 Leu Gly Gly Val Leu Ile Asp Gly Gly Lys Phe Asp Trp Thr Val Glu
235 225      230      235      240
237 Lys Asp Gly Lys Pro Val Phe Pro Tyr Phe Val Thr Pro Asp Ala Ala
238      245      250      255
240 Tyr His Gly Leu Lys Tyr Ala Asp Leu Gly Ala Pro Ala Phe Gly Leu
241      260      265      270
243 Lys Val Arg Val Gly Leu Leu Arg Asp Thr Gly Ser Thr Leu Ser Ala
244      275      280      285
246 Phe Asn Ala Trp Ala Ala Val Gln Gly Ile Asp Thr Leu Ser Leu Arg
247      290      295      300
249 Leu Glu Arg His Asn Glu Asn Ala Ile Lys Val Ala Glu Phe Leu Asn
250 305      310      315      320
252 Asn His Glu Lys Val Glu Lys Val Asn Phe Ala Gly Leu Lys Asp Ser
253      325      330      335
255 Pro Trp Tyr Ala Thr Lys Glu Lys Leu Gly Leu Lys Tyr Thr Gly Ser
256      340      345      350
258 Val Leu Thr Phe Glu Ile Lys Gly Gly Lys Asp Glu Ala Trp Ala Phe
259      355      360      365
261 Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Ile Gly Asp Val
262      370      375      380
264 Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser Asp
265 385      390      395      400
267 Glu Ala Gly Leu Ala Arg Ala Gly Val Thr Gln Ser Thr Val Arg Leu
268      405      410      415
270 Ser Val Gly Ile Glu Thr Ile Asp Asp Ile Ile Ala Asp Leu Glu Gly
271      420      425      430
273 Gly Phe Ala Ala Ile

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/746,660A

DATE: 02/12/2002

TIME: 11:32:37

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

```

274          435
277 <210> SEQ ID NO: 3
278 <211> LENGTH: 1495
279 <212> TYPE: DNA
280 <213> ORGANISM: Corynebacterium glutamicum
282 <220> FEATURE:
283 <221> NAME/KEY: CDS
284 <222> LOCATION: (287)..(1264)
286 <400> SEQUENCE: 3
287 ccattggtttc ctcagcggaa acggcttggc tatcagcact ttcacccgaa cagcctgcaa 60
289 gaagtgcgac ggctaacagg gctgggattg tectcaactt cacttcgggc tccttcttag 120
291 taataggttc gtagaaaagt ttactagcct agagagtatg cgatttcctg aactcgaaga 180
293 attgaagaat cgccggacct tgaaatggac ccggtttcca gaagacgtgc ttcctttgtg 240
295 ggttgccgaa agtgattttg gcacctgccc gcagttgaag gaagct atg gca gat 295
296                                     Met Ala Asp
297                                     1
299 gcc gtt gag cgc gag gtc ttc gga tac cca cca gat gct act ggg ttg 343
300 Ala Val Glu Arg Glu Val Phe Gly Tyr Pro Pro Asp Ala Thr Gly Leu
301      5          10          15
303 aat gat gcg ttg act gga ttc tac gag cgt cgc tat ggg ttt ggc cca 391
304 Asn Asp Ala Leu Thr Gly Phe Tyr Glu Arg Arg Tyr Gly Phe Gly Pro
305 20          25          30          35
307 aat ccg gaa agt gtt ttc gcc att ccg gat gtg gtt cgt ggc ctg aag 439
308 Asn Pro Glu Ser Val Phe Ala Ile Pro Asp Val Val Arg Gly Leu Lys
309          40          45          50
311 ctt gcc att gag cat ttc act aag cct ggt tgc gcg atc att gtg ccg 487
312 Leu Ala Ile Glu His Phe Thr Lys Pro Gly Ser Ala Ile Ile Val Pro
313          55          60          65
315 ttg cct gca tac cct cct ttc att gag ttg cct aag gtg act ggt cgt 535
316 Leu Pro Ala Tyr Pro Pro Phe Ile Glu Leu Pro Lys Val Thr Gly Arg
317      70          75          80
319 cag gcg atc tac att gat gcg cat gag tac gat ttg aag gaa att gag 583
320 Gln Ala Ile Tyr Ile Asp Ala His Glu Tyr Asp Leu Lys Glu Ile Glu
321      85          90          95
323 aag gcc ttc gct gac ggt gcg gga tca ctg ttg ttc tgc aat cca cac 631
324 Lys Ala Phe Ala Asp Gly Ala Gly Ser Leu Leu Phe Cys Asn Pro His
325 100          105          110          115
327 aac cca ctg gcc acg gtc ttt tct gaa gag tac atc cgc gag ctc acc 679
328 Asn Pro Leu Gly Thr Val Phe Ser Glu Glu Tyr Ile Arg Glu Leu Thr
329          120          125          130
331 gat att gcg gcg aag tac gat gcc cgc atc atc gtc gat gag atc cac 727
332 Asp Ile Ala Ala Lys Tyr Asp Ala Arg Ile Ile Val Asp Glu Ile His
333          135          140          145
335 gcg cca ctg gtt tat gaa ggc acc cat gtg gtt gct gct ggt gtt tct 775
336 Ala Pro Leu Val Tyr Glu Gly Thr His Val Val Ala Ala Gly Val Ser
337      150          155          160
339 gag aac gct gca aac act tgc atc acc atc acc gca act tct aag gcg 823
340 Glu Asn Ala Ala Asn Thr Cys Ile Thr Ile Thr Ala Thr Ser Lys Ala
341      165          170          175

```

→

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/746,660A

DATE: 02/12/2002

TIME: 11:32:38

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\02122002\I746660A.raw

L:5731 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67  
L:6951 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:6955 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:6956 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:6959 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:6960 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83  
L:7023 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84  
L:7026 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84